

第二章 文獻探討

本章分別簡介與本研究相關的感測方式、由 2D 相片建立 3D 臉像、從影像序列建構人體動作、用影像分析追蹤人體方式、顏色分析、與投影等相關研究。

第一節 Motion Capture System 感測方式

動作擷取系統依其感測方式可以區分為三種，茲介紹如下：

一、機械式

孫弘 (民 90)指出這是最早發展出來的動作擷取系統，利用固定於演員身上的連桿與轉軸記錄關節的旋轉角度資料，機械式最主要的缺點是不能記錄演員 Z 軸方向的位移，也就是說沒有跳躍的高度資料；優點是設備價格低廉，同時機械式因為是以裝置於演員身上的轉軸來記錄關節轉動的角度數值，因此使用場地無論是室內、室外，各種類型及範圍幾無限制，不像電磁式或光學式必須侷限在一定的區域之內，但是演員身上的連桿與轉軸等機械裝置會限制演員某些動作的表現。

二、電磁感應式

電磁感應式不同於機械式以轉軸捕捉記錄關節轉動角度或是光學式以感光點捕捉記錄關節點的位移座標資料，電磁感應式可以同時捕捉記錄關節的位移及旋轉角度資料，一般可以看到捕捉區域附近會有一至二個透明球或黑盒子，這就是磁場產生器，安裝於演員身上的感應器其實就是線圈，其原理就是利用線圈在磁場中移動會產生電流，同時因電生磁而方向性的定律，得以記錄該線圈在三度

空間中的位置及轉動角度。(孫弘，民 90)

電磁式的價格適中，稍具規模的公司均可架設，由於記錄資料齊全，因此資料轉換過程，較為正確且容易，其主要缺點是架設場地限制較大，舞台周圍必須盡量沒有感磁性金屬，由於感應器體積較小且較為貼身，因此可以捕捉如翻滾、爬行的動作。

三、光學式

光學式感測方法為採用數台(一般是 6 台以上)紅光 LED 攝影機(如圖 2-1)，演員身上黏貼表面有特殊反光塗料的光點球，經由各台攝影機拍攝到的光點位置交互運算，即可得到該關節點的空間位移，然後再經由軟體換算為關節的轉動角數據，再供動畫軟體使用。(孫弘，民 90)

光學式 Motion Capture System 的設置費用最為高昂，通常是以 LED 攝影機的數量來決定總價格，因此除非是有足夠的業務量的公司才足以負擔這樣的設施，奇特的是國內反而是以光學式系統為多數，由於是使用攝影機交互運算，因此最好架設在一個固定空間中，且由於攝影機間互相看不見及演員動作需在攝影機視角之內，最好能以適當的高度來安置攝影機，光學式是使用彈性最大的系統，因為反光點活動彈性，如大象、貓、狗之類的動物也可以貼上光點來捕捉其動作，如果架設在較大的攝影棚，也有足夠之攝影機，也可以進行多人動作同時捕捉。



圖 2-1 紅光 LED 攝影機

以上三種方法中，以光學式的對演員的動作限制較少，且是目前較多公司使用的，因此本研究也採用類以關節點辨視之方法。

丁力群 (民 87)指出在動畫軟體裡調整人物的動作一向是件不簡單的事，也因為我們眼睛所接觸到的人物動作是那麼的自然呈現，因此在製作人物動畫時，就難免會有“要求過高”的情形出現。在現實生活中，我們擺動個手臂是件輕而易舉的事，但在動畫裡，就得考慮手臂擺動時，骨架方面的連動。而本研究在使用 CCD 拍攝時，就非常要求影像的解析度和影像擷取的速度，若是解析度太底，或是影像擷取的速度不過，這都會影響最後虛擬人物的動作結果呈現。

第二節 運用二維相片組構三維臉像

為了求出演員身體關節點座標，本研究參考了(蔡榮發，民 86)二維相片組構三維臉像之模式的座標計算方式，而套用至關節點的座標計算，而本研究是使用了 3 台 CCD 各間隔 120 度的方式拍攝演員之關節點色球。

此模式運用三張右側、正面及左側同時取像的照片，以最佳化的方式分別計

算出左、右臉的控制點的三維座標，再模擬出三維臉像。用左面與正面影像來分析左臉控制點座標；用右面與正面向來分析右臉控制點座標。

在拍攝影像時有幾點必須注意：

- 一、架設三部相機時，取景需涵蓋被測者的右、正、左三面的影像。
- 二、三台相機需同時取像，所以，為達同步的目的，拍照時必須先關閉周遭光源，再把兩部相機的快門開啟，再讓第三部相機驅動閃光燈進行曝光。
- 三、設置閃光燈時，閃光燈打光範圍必須涵蓋被測者臉部各角度。
- 四、正面取像相機與兩側面取像相機間的夾角須知道，而且拍攝時，三台相機的高度與水平仰角必須相同，這些要求易於做到，因為相機之間的夾角可用一般的量角器測得，高度要維持一樣可利用一般透明水管裝水，運用連通管原理輕易的使每部相機高度都相同，水平仰角的控制則可用一般角架上所附的水平儀使得每部相機的水平仰角都相同。
- 五、被拍照者必須注意避免使頭髮覆蓋住臉部、必要時應把頭髮往後梳，鏡頭中的臉像應盡可能大。以方便後續做影像分析。

特性一：對一實體而言，無論怎樣轉動，實體上任何兩點的空間距離仍然相等。假設 a 、 b 為實體上任意相異兩點，在實體尚未轉動前之座標分別為

(x_a, y_a, z_a) 、 (x_b, y_b, z_b) ，轉動後的座標分別改變為 (x'_a, y'_a, z'_a) 、 (x'_b, y'_b, z'_b) 則

$$(x_b - x_a)^2 + (y_b - y_a)^2 + (z_b - z_a)^2 = (x'_b - x'_a)^2 + (y'_b - y'_a)^2 + (z'_b - z'_a)^2$$

可應用於構建臉像的過程中。同時取像的三張照片中的臉像，可視為臉像轉動不同角度後所拍攝的影像。故臉像上任意兩點之間的實際距離仍然不變，只是三張影像中的臉像大小可能不同，但彼此有一定的比例關係存在。

特性二：當三台相機的高度相同且仰角均為 0 度，則正面臉像中線上的任一控制點 (x, y, z) 與側面臉像中線上相對應的控制點 (x', y', z') ，其 z 值有下列的關係：

$$z' / z = \cos q$$

z : 正面臉像中線控制點的 z 值

z' : 側面臉像中線控制點的 z 值

q : 兩台相機之間的夾角

其說明如下：

圖 2-2 為影像拍攝現場示意圖。

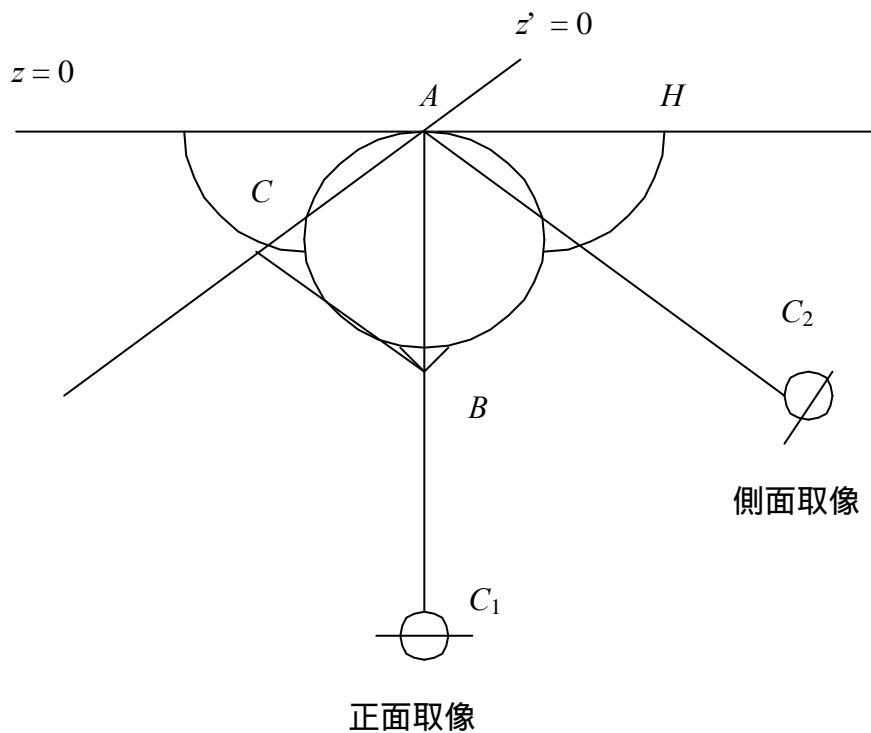


圖 2-2 影像拍攝現場示意圖

假設 B 點表正面影像的臉面中線上任意一點(如鼻尖點), C_1 、 C_2 分別表示正、側相機擺置。令 A 為正面與側面影像 z 值的基準線 \overline{AH} 又 \overline{AC} 的交點, \overline{AH} 線上任一之 z 值為 0, \overline{AC} 線上任一點之 z' 值為 0, 兩相機之間的夾角為 q 。 $z'=0$ 這條線與側面相機的鏡頭平行, 亦即與 \overline{AC} 垂直。而在正面影像中的 B 點, 令

$z = \overline{AB}$; 側面影像中的 B 點 , 令 $z' = \overline{BC}$, 且 \overline{AC} 垂直 \overline{BC} , 所以 \overline{AC} 平行 \overline{BC} , 故 $\angle ABC = \angle CAC = \theta$, 根據三角形的餘弦定理可得知 $z'/z = \cos \theta$ 。

特性三：如果三張照片是同時取像，而且三台相機高度相同，且水平仰角都為 0 度，則照片中的臉像大小有一定的比例關係，設正面與側面影像中額頭頂端(控制點 1)與下額點(控制點 4)距離分別為 s 與 s' , 其比值 $(s/s') = t$, t 為正、側面影像中臉像大小之比例。

因為從正面和側面同時取像，三張照片中臉像是處於同一個狀態，不會因為不同時間所擷取的照片而造成三張照片中的臉像表情不一或是頭部的移動而使建構臉像時產生的誤差。根據特性一、特性二、特性三，可得二維與三維座標的轉換關係式如下：

$$t \times [(x_b - x_a)^2 + (y_b - y_a)^2 + (z_b - z_a)^2] = (x'_b - x'_a)^2 + (y'_b - y'_a)^2 + (z'_b - z'_a)^2$$

$$z'/z = t \times \cos \theta$$

a 、 b 表任意相異兩控制點，其中，所有的 x 、 y 座標值皆為已知，也就是經由影像分析所測得的平面座標值。

以上為臉像之取像座標之計算方法，且是使用簡單之三角函數運算之方式，而這種方式只要加以修改，便可使用在本研究之中。

拍攝三維物體的方法一般來說可分為兩種，一種是在物體的各個角度擺設多個攝影機，另一種是只用二台攝影機，把觀測的物體放在轉盤上，轉動轉盤再用照機取像，所取得圖片就像是由另一個相機照的。有了轉盤，即使只有兩台照相機，仍能得到其他觀測點所拍到的影像。

把觀測物體放置在轉盤上，藉由轉盤的轉動，即使相機並未移動，我們仍能從不同的觀測端觀測物體，就好像有一台照相機擺在另一個觀測點。這個實際上不存在的照相機稱為虛擬照相機(林怡良，民 86)。轉盤的轉動角度可以事先控制，因此虛擬照相機的位置及方位可由實體照相機的位置及方位推得。換言之，不論虛擬照相機或實體照相機，其 3D 點及影像上 2D 點的關係，都是可以推導

得到。

本研究由於場地很大，因此採用的是第一種方式，利用多台 CCD 相機拍攝演員之動作。

林怡良(民 86)指出在找對應點之前，必須要先校正照相機。知道照相機的位置及鏡頭所擺方位，才能推出 3D 環境中的點的座標與其投影點的關係，從而到兩張圖的對應點。校正品質的好壞，影響後續的尋找對應點過程甚鉅。如果校正得到的參數不夠精準，推導 3D 及 2D 點之間的關係就會有顯著的誤差，導致無法找到兩張圖的對應點，或者找到錯誤的對應點。這個觀念在本研究中非常重要，本研究在拍攝前需將攝影機的拍攝距離及角度調整好才可以開始拍攝，否則會造成無法預期之虛擬人物動作。

林怡良(民 86)指出照相機擺設位置及方位如沒有固定，在其所得到的圖片上尋找對應點，並不是很容易。透過校正圖片，可以使圖片的對應點在同一條平行或垂直線上。校正圖片主要是把原本圖片上的點，投影到新的圖片平面。照相機的鏡心並沒有變，只是投影到的圖片平面變成一致，經過此過程得到的圖片，3D 上的點在圖片平面投影會在同一條直線上，這對縮短尋找對應點的時間幫助很大。

陳文志(民 89)指出經由影像定位的方法，可分為主動式(active)跟被動式(passive)兩類，國內外大部分研究立體視覺皆以主動式為主，如中科院發展的楊戩一號機械頭，及光學式逆向三次元量床等，其特點不外乎藉由外加雷射光線，定位建立模型。而主動式因雷射光的形式，又可分為單點式，直線式，及網狀結構式三種輔助的打光方法。而在被動式方面因為影像特徵點的搜尋不易，特徵比對速度較慢及容易受到背景及環境因素影響，故較少有被動式的研究及發展應用。另有其他小部份特殊的方法，如 CCD 自身旋轉法，資料庫比對法，或者物體固定旋轉角度融合建模 等另類方法，也是可大略估算三維座標模型的方法。

針對被動式無外加雷射的特徵搜尋法，陳文志 (民 89)提出數種被動式的方法，不須外加雷射光源，而是利用影像灰階分布為基礎，經由單張影像濾波器搜

尋方法及影像特徵比對法來找出物體的 3D 座標位置，採用不同的定位方法包括：雙 CCD 平行法、單 CCD 旋轉法、固定焦距法 等不同的立體成像法則。其中雙 CCD 平行法中又包括區塊灰階比較法、邊界法、投影法、重心法及類神經學習分類等五種方法，來定位空間中物體的三維座標。

被動式無外加雷射主要是利用兩個 CCD、影像擷取卡、二軸旋轉機台、XYZ TABLE 及電動車的硬體設備，配合 Visual Basic 及 Visual C++ 程式設計軟體做硬體界面控制，驗證所提出的數種方法，透過數種定位方法的搭配使用，將目標物體空間座標傳送至機械手臂控制器，使得機器手臂能夠準確定位到定點進行焊接修補裂縫的工作。(陳文志，民 89)

本研究使用的方法即使用被動式之拍攝方法，使用 3 台 CCD 拍攝演員身上之彩色關節點色球。

第三節 從影像序列建構虛擬實境中的人體動作

侯君昱(民 86)指出以視覺為基礎的人體外形及動作辨識在許多相關的研究領域中逐漸獲得重視。在這些研究課題裡，主要的目的有二：重建立體的人體結構和分析人體的動作序列。

侯君昱(民 86)提出了一個快速的解決方法，它可從現有的影像序列取出其中的人體外形，並在虛擬環境中建構對應的三維人體影像。此方法依序可分為三個主要的階段：取得影像中最大的移動區域、模型比對以及深度圖的建立。

首先，第一個階段，計算影像序列裡所有像框的像素值，以取得固定的背景影像。接下來將原始影像消去背景後，即可得到包括雜訊在內的移動區域。利用一個取得最大移動區域的演算法，包含人體的一塊區域可被快速地抽取出來。在

第二個階段中，我們以事先建立好的人體模型來比對這個被抽取出的影像，其辨認順序依次為頭部、軀幹、大腿、肩膀及四肢，並且取得它們之間的相對位置。在最後一個階段裡，以肢體的角度和長度來建立人體在三維空間中的深度圖。

上述之研究方法為先去背景，再做影像比對的動作，而本研究發現在去背景的過程中，若背景有和演員身上色球相同顏色之色塊，則在辨視時便會無法將背景去除，而造成多出關節點的現象。因此本研究為了減少關節點之判別錯誤，因此限制背景使用黑色布幕，而演員必須穿著黑色緊身衣服。而在動作辨視的部份，本研究則是參考了光學式之方法拍攝色球關節點。

第四節 利用視訊計算技術來建構虛擬人的 HIVE 編解碼

在民國 89 年，林中青發展了一套即時的 HIVE 系統，人體動作分析合成系統，這套系統是由一台固定的攝影機拍攝在真實環境下動作的人體，再利用視訊計算技術來分析及追蹤此人體，並即時地在虛擬環境下產生一個虛擬人來模擬真人的動作。

林中青(民 89)提出一套模型基礎的方法來分析真人的動作，然後合成一個虛擬人來模擬真人的動作。這個系統包含四個階段：背景分離、建立三維人體模型、人體動作分析及合成虛擬人。在背景分離階段，為了要在複雜的背景中擷取出完整的前景物件，必須先拍攝一段靜態的背景影像序列，統計出背景模型參數以建立背景模型。然後從影像序列中去除掉與背景模型相似的像點以擷取出前景物件。在建立三維人體模型階段，分析前景物件影像，然後計算出三維人體模型的形狀參數，如：軀幹、頭部、手部及腿部的長度及半徑，以建立三維人體模型。

在人體動作分析階段，改變三維人體模型的運動參數(各關節的角度)產生數個三維人體模型，並將這些三維人體模型投影在二維平面上，然後將三維人體模型的二維投影和前景人體物件作匹配比較，找出最相似的人體模型，再由此人體模型所對應的各關節的角度我們即可找出人體模型的運動參數。最後將所有的運動參數組合起來，利用三維電腦繪圖技術(OpenGL)產生一個三維的虛擬人來模擬真人的動作。

使用上述之方法時，若每次演員不同，體型不同，就必須再度測量該演員之體型的各項參數，較不方便。

第五節 顏色分析觀念，彩色原理

光是一種電磁波幅射能，而電磁波之波長範圍在人眼所能觀察到的波長只有一小範圍，可見光波長從 380 nm 到 780 nm 之間，通常這個範圍的頻譜稱為可見光譜(Visible spectrum)。CIE 制定紅光波長為 $R(\lambda)=700$ nm，綠光波長為 $R(\lambda)=546.1$ nm，藍光波長為 $R(\lambda)=435.8$ nm。而在數位影像中以數字如(0~255)來表示強度。(謝振榆、傅毓芬，民 88)

每一種色彩都是光波特別混合的一種結果，而人類的視覺官感，建立在三原色的基本理論上。三原色的等量混合可以產生白光，非等量混合，可以產生一千六百多萬種顏色中的任何一種顏色。(謝振榆、傅毓芬，民 88)

由於光波的不同，所以產生的色彩也不一樣。光的種類很多，和色彩變化有密切關係的只有兩種，那就是直接光和反射光。直接光是可以從光源直接看到的光，如日光、燈光、電腦顯示器發出的光。直接光的原色是增原色把 R、G、B 中任兩種顏色，採取等量混合，會產生一種補色。如紅與藍等量混合會產生紫色，紅與綠等量混合，會產生一種補色，藍與綠等量混合會產生青色。若把紅、藍及綠等量混合會產生白光。(謝振榆、傅毓芬，民 88)

反射光是直接光碰到物體所產生的一種反射現象。在實際上物體的本身並沒有顏色，是物體的表面吸收了光波的某些部份，而又反射了光波的某些部分，才有了顏色。例如在黑夜中任何顏色都看不見，是因為沒有光，沒有反射，才沒有顏色，所以無法見到顏色。(謝振榆、傅毓芬，民 88)

反射光作用是在減色原理上，青(C)、紫(M)、黃(Y)是減原色，是加原色的補色。而紅、藍、綠是從減三原色產生的。例如，紫與黃等量混合產生紅色，黃與青等量混合產生綠色，青與紫等量混合產生藍色。把紫、表黃三色等量混合，會產生黑即是無色。而無色的黑並不是黑色，因此在印刷顏色系統上除了減三原色之外，還要增加天然黑，即四色印刷的理由。

本研究是利用反射光之原理，因此光線的控制非常的重要。